

<http://www19.ipdl.ncipi.go.jp/PA1/result/detail/main/wAAaIsaGs0DA407036484P...> 2005/05/17

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-36484

(43)公開日 平成7年(1995)2月7日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 7/02		8946-5H		
9/14	F	8946-5H		
H 0 3 H 17/02	E	8842-5 J		
	L	8842-5 J		

審査請求 未請求 請求項の数3 OL (全 9 頁)

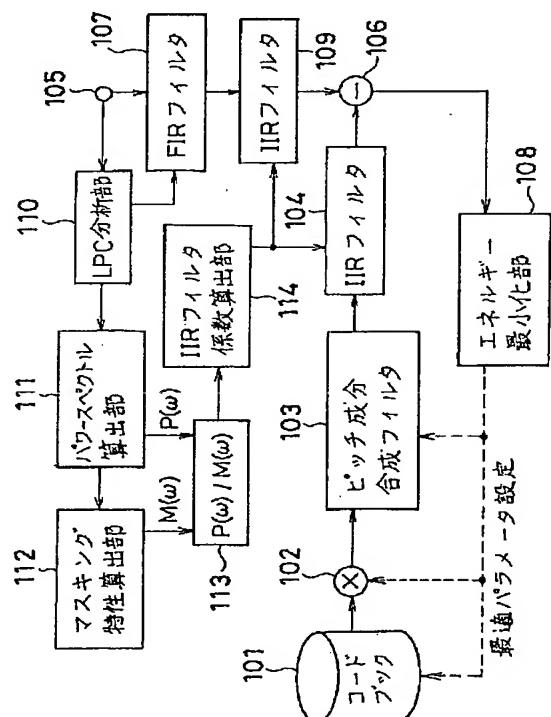
(21)出願番号	特願平5-180380	(71)出願人	000005049 シャープ株式会社 大阪府大阪市阿倍野区長池町22番22号
(22)出願日	平成5年(1993)7月21日	(72)発明者	森尾 智一 大阪府大阪市阿倍野区長池町22番22号 シャープ株式会社内
		(74)代理人	弁理士 川口 義雄 (外1名)

(54) 【発明の名称】 音響信号符号化装置

(57) 【要約】

【目的】 音響信号を圧縮符号化する装置において、符号化によって生じる量子化ノイズを、聴覚マスキング特性を利用して聞こえにくくすることで、音声品質を向上する、高圧縮で処理量の少ない音響信号符号化装置を提供する。

【構成】音響信号のパワースペクトルを求めるパワースペクトル算出部(111)と、聴覚マスキングスペクトル特性を求めるマスキング特性算出部(112)と、音響信号の逆パワースペクトル特性を有する有限インパルス応答フィルタ(107)と、音響信号のパワースペクトル特性を聴覚マスキングスペクトル特性で除したスペクトル特性を有する全極形無限インパルス応答フィルタ(104、109)を備えており、有限インパルス応答フィルタと全極形無限インパルス応答フィルタで聴覚的重み付け処理を行う。



【特許請求の範囲】

【請求項 1】 音響信号のパワースペクトルを求める手段と、聴覚マスキングスペクトル特性を求める手段と、該音響信号の逆パワースペクトル特性を有する第 1 フィルタリング手段と、該音響信号のパワースペクトル特性を該聴覚マスキングスペクトル特性で除したスペクトル特性を有する第 2 フィルタリング手段とを備えており、該第 1 フィルタリング手段及び該第 2 フィルタリング手段により聴覚的重み付け処理を行うことを特徴とする音響信号符号化装置。

【請求項 2】 前記音響信号のパワースペクトルから自己相関系列を求める逆フーリエ変換手段と、自己相関系列から前記第 2 フィルタリング手段の係数を算出する手段を更に備えていることを特徴とする請求項 1 に記載の音響信号符号化装置。

【請求項 3】 対数パワースペクトルを求める手段と、該対数パワースペクトルから逆フーリエ変換によってケプストラムを求める手段と、該ケプストラムから前記第 2 フィルタリング手段の係数を算出する手段とを更に備えていることを特徴とする請求項 1 に記載の音響信号符号化装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、オーディオ信号や音声信号を圧縮符号化して通信または蓄積する音響信号符号化装置に関する。

【0002】

【従来の技術】 第 1 の従来技術として、音声信号を圧縮符号化する際に、符号化で生じる量子化ノイズを、聴覚マスキング特性を利用して、スペクトルシェイピングする技術がある。その一例としては、“A New Model of LPC Excitation for Producing Natural-Sounding Speech at Low Bit Rates”, B. S. Atal and J. R. Remde, IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 614-617, 1982、が知られている。

【0003】 これは音声信号を線形予測分析して得られる線形予測係数を用いて、数 1 で表される伝達特性を持つフィルタで、量子化誤差波形をフィルタリングし、そのフィルタリングされた誤差波形のエネルギーを最小化するように符号化処理を行う手法である。

【0004】

【数 1】

$$W(z) = \frac{1 - \sum_{k=1}^p \beta^k a_k z^{-k}}{1 - \sum_{k=1}^p \gamma^k a_k z^{-k}} \quad (1)$$

【0005】 上記式 (1) において、 a_k は k 次の線形予測係数、 p は予測次数、 β 、 γ は $0 \leq \gamma \leq \beta \leq 1$ の定

数をそれぞれ表す。

【0006】 この聴覚的重み付けフィルタを用いた音声符号化方式である、符号帳励振線形予測符号化 (Code-Excited Linear Predictive Coding、以後 CELP と記す) は、例えば、“Code-Excited Linear Prediction (CELP): High-Quality Speech at Very Low Bit Rates”, M. R. Schroeder and B. S. Atal, IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 937-940, 1985 に示されている。図 4 はその構成を示すブロック図である。

【0007】 図 4 において、 $1/A(z)$ は、式 (2) で表される音声の線形予測合成フィルタである。

【0008】

【数 2】

$$\frac{1}{A(z)} = \frac{1}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (2)$$

【0009】 上記式 (1) において $\gamma = 0$ 、 $\beta = 1$ に設定して、上記式 (2) で表される音声の線形予測合成フィルタとこの聴覚的重み付けフィルタを合成すると、式 (3) のように簡略化される。

【0010】 この場合、図 4 のブロック図は、図 5 に示す構成に変更される。

【0011】

【数 3】

$$H(z) = \frac{1}{1 - \sum_{k=1}^p \gamma^k a_k z^{-k}} \quad (3)$$

【0012】 上述した従来の技術では聴覚的重み付けフィルタは、聴覚マスキング特性を非常に簡単に近似した特性で表している。

【0013】 第 2 の従来技術として、オーディオ信号の圧縮符号化で用いられている技術がある。この方式は、第 1 の従来技術より積極的に聴覚マスキング特性を利用している。

【0014】 図 6 に、MPEG で用いられている音響信号の符号化部の動作シーケンスを示す。その一例は、

「音響信号の高効率符号化—MPEG オーディオ符号化方式」後藤、日本音響学会誌 47 巻 12 号 pp. 966-969、1991 に示されている。

【0015】 図 6 のフローの右上において、入力信号を FFT を用いパワースペクトルを求め、パワースペクトルの情報等から、聴覚マスキング特性を算出している。MPEG Layer 1、2 では、基本的には帯域分割符号化を用いており、マスキング特性の情報等から、各帯域毎の符号化ビットを決定している。

【0016】第3の従来技術としては、第1と第2の技術を融合した技術がある。パワースペクトル情報から聴覚マスキング特性を求め、その逆特性を持つ聴覚的重み付けフィルタを用い、量子化誤差波形のエネルギーを最小化するように符号化処理を行う手法である。その一例は、“Some Experiments in Perceptual Masking of Quantizing Noise in Analysis-By-Synthesis Speech Codes”, R. Drogo De Iacovo and R. Montagna, EUROSPEECH, p. 825-828, 1991に示されている。

【0017】この方式においては、ヒルベルト変換の技術を用いて、聴覚マスキング特性のパワースペクトル特性を持つ、最小位相有限インパルス応答フィルタ（以後FIRフィルタと記す）を設計し、その逆フィルタを聴覚的重み付けフィルタとして使用している。

【0018】

【発明が解決しようとする課題】しかしながら、上述した第1の従来技術における聴覚的重み付けフィルタの特性は、簡単な近似によって求められているので人間の聴覚マスキング特性とは異なっており、量子化ノイズを十分に隠蔽することができないという問題点があった。

【0019】また、上述した第2の従来技術においては、マスキング特性は、人間の聴覚マスキング特性のモデルに従って求めてはいるが、最終的に帯域分割符号化を用いており、ビット配分等の付加情報も必要で、圧縮率が十分に低くできないという問題点があった。

【0020】更に、上述した第3の従来技術においては、上記2つの問題点に対処し、聴覚マスキング特性を考慮し、聴覚的重み付けフィルタを用いることで、圧縮率の高い符号化方式が実現できる。しかしながら聴覚的重み付けフィルタはFIRフィルタで構成されているゆえ、同一フィルタ次数で振幅周波数特性を近似する観点からは無限インパルス応答フィルタ（以後IIRフィルタと記す）より劣るという課題と、第1の従来技術で説明したような、聴覚的重み付けフィルタと、音声の線形予測合成フィルタとの合成処理による処理の簡易化が困難であるという問題点があった。

【0021】本発明の目的は、上述した従来技術における問題点に鑑み、量子化ノイズを十分に隠蔽でき、圧縮率が十分に低くできると共に全体の処理を簡易化できる音響信号符号化装置を提供することにある。

【0022】

【課題を解決するための手段】本発明の目的は、音響信号のパワースペクトルを求める手段と、聴覚マスキングスペクトル特性を求める手段と、音響信号の逆パワースペクトル特性を有する第1フィルタリング手段と、音響信号のパワースペクトル特性を聴覚マスキングスペクトル特性で除したスペクトル特性を有する第2フィルタリング手段とを備えており、第1フィルタリング手段及び第2フィルタリング手段により聴覚的重み付け処理を行う音響信号符号化装置によって達成される。

【0023】本発明の音響信号符号化装置は、音響信号のパワースペクトルから自己相関係列を求める逆フーリエ変換手段と、自己相関係列から第2フィルタリング手段の係数を算出する手段を備えるように構成されてもよい。

【0024】本発明の音響信号符号化装置は、対数パワースペクトルを求める手段と、対数パワースペクトルから逆フーリエ変換によってケプストラムを求める手段と、ケプストラムから第2フィルタリング手段の係数を算出する手段とを備えるように構成されてもよい。

【0025】

【作用】本発明の音響信号符号化装置では、音響信号のパワースペクトルを求め、聴覚マスキングスペクトル特性を求め、第1フィルタリング手段は音響信号の逆パワースペクトル特性を有し、第2フィルタリング手段は音響信号のパワースペクトル特性を聴覚マスキングスペクトル特性で除したスペクトル特性を有し、第1フィルタリング手段及び第2フィルタリング手段により聴覚的重み付け処理を行う。

【0026】本発明の音響信号符号化装置では、逆フーリエ変換手段は音響信号のパワースペクトルから自己相関係列を求め、自己相関係列から第2フィルタリング手段の係数を算出する。

【0027】本発明の音響信号符号化装置では、対数パワースペクトルを求め、対数パワースペクトルから逆フーリエ変換によってケプストラムを求め、ケプストラムから第2フィルタリング手段の係数を算出する。

【0028】

【実施例】以下、図面を参照して本発明の音響信号符号化装置の実施例を説明する。

【0029】図1は、本発明の音響信号符号化装置の第1実施例の構成を示すブロック図であり、CELPシステムを用いた例を示す。

【0030】図1の音響信号符号化装置は、音響信号の入力端子105、入力端子105に接続されており音響信号を線形予測分析（以後LPC分析と記す）するLPC分析部110、LPC分析部110に接続されておりLPC分析結果から信号のパワースペクトル $P(\omega)$ を算出するパワースペクトル算出部111、パワースペクトル算出部111に接続されており信号のパワースペクトルからマスキング特性 $M(\omega)$ を算出するマスキング特性算出部112、パワースペクトル算出部111及びマスキング特性算出部112に接続されており信号のパワースペクトルをマスキング特性で割算する割算器113、割算器113に接続されており割算器113で求めたスペクトル比特性からIIRフィルタ係数を求めるIIRフィルタ係数算出部114、入力端子105及びLPC分析部110に接続されており入力信号を聴覚的重み付けするための第1フィルタリング手段であるFIRフィルタ107、FIRフィルタ107及びIIRフ

フィルタ係数算出部 114 に接続されており入力信号を聴覚的重み付けするための第 2 フィルタリング手段の一部である IIR フィルタ 109、CELP 音声符号化の励振符号帳（コードブック）101、コードブック 101 に接続されており励振信号を増幅する増幅部 102、増幅部 102 に接続されておりピッチ成分を合成するピッチ成分合成フィルタ 103、ピッチ成分合成フィルタ 103 及び IIR フィルタ係数算出部 114 に接続されており音声スペクトル合成フィルタと聴覚的重み付けフィルタを合成した特性を持つ第 2 フィルタリング手段の他の一部である IIR フィルタ 104、IIR フィルタ 104、109 に接続されており聴覚的重み付けされた入力信号と聴覚的重み付けされた再生信号の差分をとる減算部 106、減算部 106 に接続されており差分波形のエネルギーを最小化するように符号化パラメータを設定するエネルギー最小化部 108 によって構成されている。

$$P(\omega) = \frac{1}{\sqrt{\left(1 - \sum_{k=1}^p a_k e^{-j\omega k}\right) \left(1 - \sum_{k=1}^p a_k e^{j\omega k}\right)}} \quad (4)$$

【0034】上記式（4）において、 $\omega = 2\pi F_s$ で F_s はサンプリング周波数である。

【0035】上述の説明では、LPC 分析の結果から入力信号のパワースペクトルを算出したが、入力信号をフーリエ変換して算出してもよい。この場合、LPC スペクトルより周波数分解能を高く求められるので、マスキング特性の算出がより精度よく計算できる。

【0036】マスキング特性算出部 112 は、入力信号のパワースペクトルから、マスキングスペクトル特性を算出する。本処理手順の概要は、パワースペクトルを聴覚の臨界帯域幅毎に分解し、全ての臨界帯域毎に、入力信号による量子化雑音のマスキング曲線を算出し、信号帯域全体に渡る最小可聴値及び時間軸でのマスキング等を考慮してマスキング曲線 $M(\omega)$ を算出する。マスキ

$$\begin{aligned} W(z) &= \frac{1}{M(z)} = A(z) F(z) \\ F(z) &= \frac{1}{A(z)} \frac{1}{M(z)} \end{aligned} \quad (5)$$

ここで $z = e^{-j\omega}$

【0039】フィルタ $F(z)$ の振巾伝達特性は、入力信号のパワースペクトル $P(\omega)$ を、マスキングスペクトル $M(\omega)$ で除した伝達特性を持つフィルタと考えることができる。

【0040】このフィルタ $F(z)$ を全極形 IIR フィ

【0031】本実施例では、聴覚的重み付けフィルタの構成法が上述した図 5 の従来技術と異なる。以下では聴覚的重み付けフィルタの構成法に重点をおいて説明する。

【0032】入力端子 105 から入力した信号は、ある一定の時間長毎に区分化処理される。これをフレームと呼ぶことにする。1 フレームの信号は LPC 分析部 110 で線形予測係数が算出される。この線形予測係数は上記式（1）で示す聴覚的重み付けフィルタの分子項である FIR フィルタ 107 の係数として設定される（但し以後、上記式（1）において、 $\beta = 1$ とする）。算出された線形予測係数から振巾伝達特性をパワースペクトル算出部 111 で計算する。上記式（2）で表された伝達特性から下記に示す式（4）でパワースペクトルが算出される。

【0033】

【数 4】

ング曲線の算出は、種々提案されており、その一例としては“Estimation of Perceptual Entropy Using Noise Masking Criteria”, J. D Johnston, IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 2524-2527, 1988 がある。

【0037】符号化による量子化ノイズは、このマスキング曲線の形状に従ってシェイピングされるように聴覚的重み付けフィルタを設計する。即ち、聴覚的重み付けフィルタのパワースペクトルは、マスキングスペクトルの逆特性を持つ必要がある。ここで下記の式（5）に示す関係が成り立つフィルタ $F(z)$ を考える。

【0038】

【数 5】

ルタで実現した場合には、聴覚的重み付けフィルタ $W(z)$ と音声合成フィルタ $1/A(z)$ を合成すると、下記に示す式（6）のように簡略化できる。

【0041】

【数 6】

$$H(z) = \frac{1}{A(z)} W(z) = \frac{1}{A(z)} A(z) F(z) = F(z) \quad (6)$$

【0042】上記動作を行うために、割算部113で $P(\omega)/M(\omega)$ を求め、IIRフィルタ係数算出部114で、 $P(\omega)/M(\omega)$ で示されるパワースペクトルからIIRフィルタ係数を算出する。

【0043】ここで、フィルタ $F(z)$ の伝達関数を式(7)に示す。

【0044】

【数7】

$$F(z) = \frac{1}{1 - \sum_{k=1}^q f_k z^{-k}} \quad (7)$$

【0045】式(7)において、 q はIIRフィルタの次数で、音声の線形予測次数と一致している必要はない。 f_k はIIRフィルタ係数算出部114で算出されたIIRフィルタの k 次の係数である。

【0046】以上の処理で、LPC分析部110で求めた線形予測係数が設定された $A(z)$ の伝達関数を持つFIRフィルタ107と上述したIIRフィルタで、入力信号を聴覚的重み付け処理する。また、IIRフィルタ109と同じ係数が設定されたIIRフィルタ104によって、聴覚的重み付けされた再生信号を得る。この後の符号化処理は、一般のCELP符号化方式と同じであり、概略だけ説明すると、聴覚的重み付けされた入

力信号と、聴覚的重み付けされた再生信号の、誤差エネルギーが最小になるように、符号化のパラメータを決定する。

【0047】次に、パワースペクトル $P(\omega)$ と、マスキングスペクトル $M(\omega)$ から、IIRフィルタ係数を算出する処理法を説明する。

【0048】図2は、逆フーリエ変換と正規方程式を解くことにより、パワースペクトル $P(\omega)$ とマスキングスペクトル $M(\omega)$ から、IIRフィルタ係数を算出する処理手順を示す。

【0049】以下、図2を参照して説明する。

【0050】まず、下記の式(8)に示すようにパワースペクトルを定義する。

【0051】

【数8】

$$S(\omega) = \frac{P(\omega)}{M(\omega)} \quad (8)$$

【0052】パワースペクトル $S(\omega)$ と自己相関関数 $R(\tau)$ の間には、下記の式(9)で示すような関係があるので、 $\tau=0 \sim q$ の範囲で、FFTの手法等を使い、自己相関系列を算出する。

【0053】

【数9】

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega\tau} d\omega \quad (9)$$

【0054】次に自己相関係数から、IIRフィルタ係数への変換は、音声の線形予測分析で一般的に用いられているように、式(10)の正規方程式を解くことで求

められる。

【0055】

【数10】

$$\begin{pmatrix} R(0) & R(1) & R(2) & \dots & R(q-1) \\ R(1) & R(0) & R(1) & \dots & R(q-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R(q-1) & R(q-2) & R(q-3) & \dots & R(0) \end{pmatrix} (f_1 \ f_2 \ \dots \ f_q)^T = -(R(1) \ R(2) \ \dots \ R(q))^T \quad (10)$$

【0056】式(10)において、 $(\dots)^T$ は行列の転置操作を表す。

【0057】上述した操作で、IIRフィルタの係数が算出される。

【0058】図3は、準同形処理によるケプストラムを用いてIIRフィルタ係数を算出する他の手法を示す。

【0059】ここではパワースペクトル $P(\omega)$ を、マ

スキングスペクトル $M(\omega)$ で除す演算を、対数領域で行うので、式(11)に示す処理を行う。これは図3では、 $P(\omega)$ 、 $M(\omega)$ をそれぞれ対数演算部301、302で対数化して、演算部303で減算することに相当する。

【0060】

【数11】

$$\text{Log } S(\omega) = \text{Log } P(\omega) - \text{Log } M(\omega) \quad (11)$$

【0061】この $\text{Log } S(\omega)$ を逆FFT演算部304で逆フーリエ変換すると、式(12)によって、ケプストラム c_n が算出される(「音声情報処理の基礎」

$$C_n = \frac{1}{2\pi} \int_0^{2\pi} \text{Log } S(\omega) e^{jn\omega} d\omega \quad (12)$$

【0063】ケプストラム c_n の低次部分がスペクトル構造を表しているので、ケプストラム窓(例えば、 $w_n = 1 : n = 1 \sim q$, $w_n = 0 : n > q$)で窓掛けする。こうして求めたケプストラム c_n から、式(13)によってIIRフィルタの係数が算出される。

【0064】

【数13】

$$f_k = C_k + \sum_{m=1}^{k-1} \frac{m}{n} C_m f_{k-m} \quad (13)$$

【0065】ただし、式(13)において、 k は $k=1 \sim q$ の整数である。

【0066】以上の説明はCELPシステムで説明したが、マルチパルス符号化等の聴覚的重み付けフィルタを構成要素として持つシステムへも容易に応用できる。

【0067】なお、符号化装置の符号化に関する部分を変更することにより、他の部分は一切変更なしで復号装置を実現できる。

【0068】

【発明の効果】本発明の音響信号符号化装置は、音響信号のパワースペクトルを求める手段と、聴覚マスキングスペクトル特性を求める手段と、音響信号の逆パワースペクトル特性を有する第1フィルタリング手段と、音響信号のパワースペクトル特性を聴覚マスキングスペクトル特性で除したスペクトル特性を有する第2フィルタリング手段とを備えており、第1フィルタリング手段及び第2フィルタリング手段により聴覚的重み付け処理を行うので、符号化で生じる量子化ノイズを、聴覚的重み付けフィルタによって、ノイズシェイピングすることができ、人間の聴覚特性の聴覚マスキングを利用することにより、雑音を聞こえにくくして再生音質を向上できる。また、聴覚的重み付けフィルタは、音声の線形予測合成フィルタと合成することで、簡易化することができ、符号化演算量を削減することができる。

【0069】本発明の音響信号符号化装置は、音響信号のパワースペクトルから自己相関系列を求める逆フーリエ変換手段と、自己相関系列から第2フィルタリング手段の係数を効果的に算出することができる。

【0070】本発明の音響信号符号化装置は、対数パワ

ー、中田、オーム社、pp. 99-103、参照)。

【0062】

【数12】

ースペクトルを求める手段と、対数パワースペクトルから逆フーリエ変換によってケプストラムを求める手段によって、ケプストラムから第2フィルタリング手段の係数を効果的に算出することができる。

【図面の簡単な説明】

【図1】本発明の音響信号符号化装置の一実施例の構成を示すブロック図である。

【図2】本発明の音響信号符号化装置におけるフィルタ係数算出の一例を説明するためのフローチャートである。

【図3】本発明の音響信号符号化装置におけるフィルタ係数算出の他の一例を説明するためのブロック図である。

【図4】従来技術のCELP音声符号化方式を説明するためのブロック図である。

【図5】従来技術のCELP音声符号化方式の、聴覚的重み付けフィルタ処理の簡易化を説明するためのブロック図である。

【図6】従来技術の聴覚マスキングを考慮した符号化方式を説明するためのフローチャートである。

【符号の説明】

101 CELPの励振符号帳(コードブック)

102 掛算部

103 ピッチ成分合成フィルタ

104, 109 全極形IIRフィルタ

105 入力端子

106 減算部

107 FIRフィルタ

108 誤差エネルギー最小化部

110 線形予測分析部

111 パワースペクトル算出部

112 マスキング特性算出部

113 スペクトル比算出部

114 全極形IIRフィルタ算出部

301, 302 対数演算部

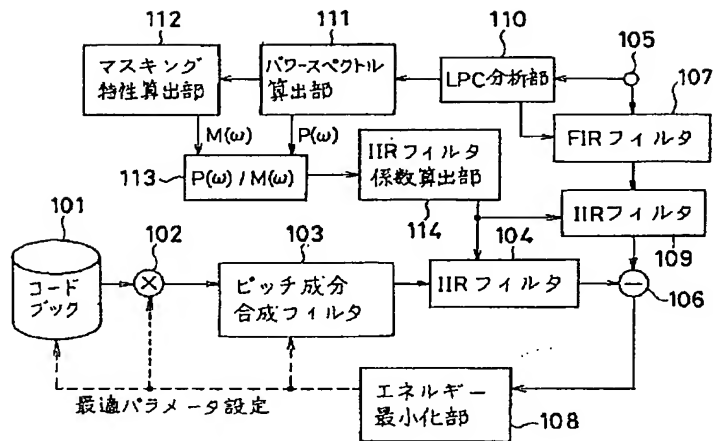
303 減算部

304 逆フーリエ変換部

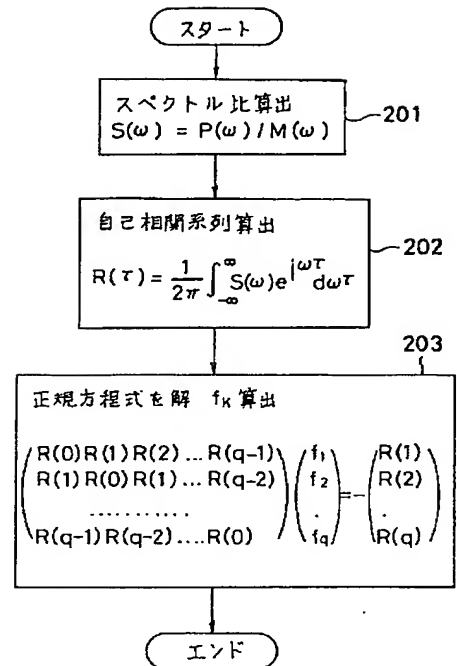
305 ケプストラム窓掛け部

306 ケプストラムから予測係数への変換部

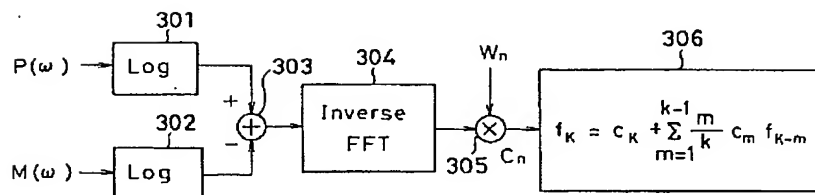
【図 1】



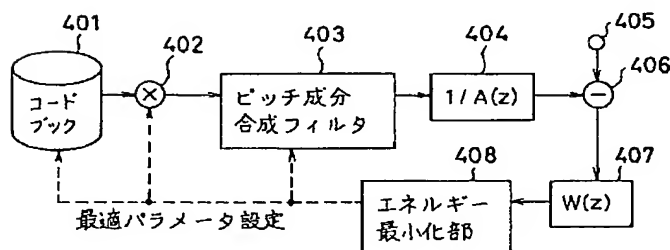
【図 2】



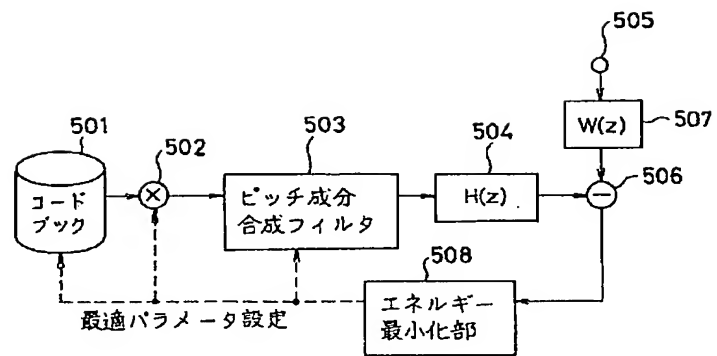
【図 3】



【図 4】



【図 5】



【図6】

